Author's response to reviews

Title: The effect of antenatal education in small classes on obstetric and psycho-social outcomes - a systematic review

Authors:

Carina S Brixval (Cabr@niph.dk)
Solveig F Axelsen (Sfax@niph.dk)
Stine G Lauemøller (Sgl@niph.dk)
Stig K Andersen (Skan@niph.dk)
Pernille Due (Pdu@niph.dk)
Vibeke Koushede (Vibe@niph.dk)

Version: 3
Date: 26 November 2014

Author's response to reviews: see over
Dear Adrienne Stevens,

Thank you for giving us the opportunity to resubmit our manuscript ‘The effect of antenatal education in small classes on obstetric and psycho-social outcomes – a systematic review’ (ref.nr 5817652331359375) for publication in Systematic Reviews. We appreciate the suggestions from the handling editor and the reviewers and believe the revisions have strengthened the manuscript.

We have addressed your and each of the reviewers’ suggested revisions below. All revisions in the manuscript are highlighted with 'track changes'. We have also added a comment about which reviewer raised the issue and which question the reply is an answer to.

In addition to the revisions suggested by reviewers we have revised the additional file 3_PRISMA flow diagram since no meta-analysis was conducted.

Yours sincerely,

Carina Sjöberg Brixval, MSc Public health
National Institute of Public Health, University of Southern Denmark
Email: cabr@niph.dk
Addr. Øster Farimagsgade 5A, 2., DK-1353 København K, Denmark
Handling Editor's comments:

1. Scope of relevant intervention. Can authors provide some parameters around how they would define 'small'? This should be stated in the 'Experimental and control conditions' section on page 5. On page 16 (lines 21-22), authors state that the 'type and arrangement of antenatal education' was the focus. It is not to me what this means; can examples be provided?

Reply 1: Thank you for taking the time to review our manuscript and for your valuable comments. We have added to the 'Experimental and control conditions' section that below 20 participants is considered a small class. This is by service providers considered a class size where interaction between the participants as well as with the educator is possible. Also, in the literature this is the typical upper limit for class size in this kind of intervention.

Type and arrangement of antenatal education means that we are interested in investigating the effect of participating in antenatal education in small classes as they provide the possibility of interaction among participants and with the educator. We have moved this paragraph to the section: ‘strengths and limitations’ and expanded the sentence.

2. Scope of relevant comparators. I am confused as to what is deemed relevant by authors in terms of 'other types of education classes' on page 5. On page 9, lines 9-10, authors state that one comparison group is 'individual prenatal care'. Isn't this a 'no intervention' comparison? Authors need to provide one or more examples of what they mean by 'other types of education classes'. It is clear from the PRISMA flow diagram that individual education is not relevant.

Reply 2a: We agree that this aspect needs clarification. When conducting the review we aimed at including all trials evaluating the effect of small classes and therefore we chose a broad comparison group definition. An example of ‘other types of educational programs’ could be e.g. pregnancy exercise or group programs with a smaller intervention dose than the experimental condition. We have included an example in the manuscript.

We agree that ‘no intervention’ is the same as ‘treatment as usual’ and have rephrased this to ‘standard care’ throughout the manuscript.

On pg 16, lines 22 and 23, authors state that two programs with the same dose of antenatal education in small classes were excluded. Is the problem that both arms had small classes? Wouldn't a second program constitute another type of education class? Please explain.

Reply 2b: We are interested in examining the influence of small classes compared to other types of antenatal education. Therefore we have excluded trials where both the experimental group and the control group received education in small classes and where the only difference was the content. We have rephrased this section in the manuscript.

On pg 5, it would be helpful to give the reader an example of what is meant by ‘treatment as usual’ (e.g., antenatal care only).
Might it be that 'no intervention' and 'treatment as usual' are the functionally the same category given that all women would continue to receive prenatal care and interventions including or with only men as participants wouldn't receive any education, as usual?

Reply 2c: We have now included the proposed example of 'treatment as usual'.

Thank you for pointing to this issue of 'no intervention' vs. 'treatment as usual'. We agree that in a Western setting all women are offered individual pregnancy care and therefore the two control conditions are functionally the same. We have therefore decided to delete the comparison group 'no intervention'.

3. Lumping of interventions for analysis. One reviewer flagged the issue of the comparability of the interventions. I would agree to this assessment, and I think that the issue at hand is the lumping of interventions in the analysis. Lumping all interventions together seems too heterogeneous with different intervention facets and different populations (in some cases) to make sense of considering all of it together. I suggest to the authors to conduct separate analyses according to intervention. Intervention groupings could look something like: depression-prevention, breastfeeding, psychosocial, general antenatal care, psychoeducational, couple-focused, and self-hypnosis. (However, I am not a content expert, and my suggested groupings may need adjustment according to content expertise.)

Within each intervention grouping, though, the results need to be presented separately according to comparison type. For example (hypothetical), results from breastfeeding education compared with no intervention would be reported separately (i.e., in a separate section) from breastfeeding education compared with another education program. The results from the two different studies would not be lumped together. I will need to see a revised version of the presentation of results to determine whether I agree with the authors' interpretation.

Reply 3: Thank you very much for your suggestion and help considering the best way to present data. We have now made the analysis separately for each of the intervention groupings. Due to the diverse content of both the experimental and control conditions this has resulted in 12 different comparisons. We have rewritten the 'effects of interventions' section in the results section and also the first part of the discussion section. Likewise additional files 6 and 7 have been changed according to your suggestions. These parts have not been “track-changed” as they have been completely re-written.

4. Reporting details of comparisons in studies. Although authors state that most studies use the comparison group of 'standard care', they do not provide the details in the Characteristics of studies table for many of the studies. Specifying the details of the control group is essential for the reader to know and to understand how it fits the eligibility criteria. As discussed above, authors will also need to consider the similarity of the control arm to determine which studies should be analyzed together.

Reply 4: We agree that this is valuable information to the reader. We have now incorporated this information into the additional file 5; characteristics of included studies table. In this table, we have also
made a few corrections in the text making the description of experimental and control conditions more similar between trials (additional file 5).

5. Risk of bias assessment. I disagree with the authors' decision to exclude the assessment of performance bias. A study's inability to blind intervention deliverers and/or recipients does not negate the influence this bias may have on study results. I request of authors to provide such assessments and to incorporate them in their overall risk of bias assessments.

Further, current guidance and best practice indicates that some assessments should be made per outcome, i.e., those addressing detection (outcome assessor) and attrition (incomplete outcome data) biases. For detection bias, which outcomes are subjective versus objective? For which outcomes did patients provide a response to that might have been biased by their knowledge of the intervention they received, etc? For attrition bias, if items that one would assess are the same for all outcomes, this needs to be stated.

Sample size is not an issue of bias, but relates to imprecision of effect estimates.

Reply 5: Thank you. We have included the assessment of bias related to blinding of participants and educators. Due to the difficulty of blinding participants and educators in these kind of educational trials, we have extended the ‘risk of bias’ assessment tool and included the following in the methods section: “Due to the type of intervention we expected a high level of bias for the domain ‘blinding of participants and educators’ as it is most likely not possible to blind participants and educators. We therefore also categorized trials as overall ‘lower risk of bias’. If a trial was rated ‘low risk of bias’ in all the risk of bias domains listed above except ‘blinding of participants and educators’, the trial was categorized as overall ‘lower risk of bias’.” This has now been explained in the ‘risk of bias assessment’ section in the Method section.

Likewise, we have incorporated your second suggestion of reporting risk of bias separately for subjective and objectively measured outcomes. This has been incorporated into the methods-section and the “Risk of bias in included studies” in the results section as well as the ‘Quality of the evidence’ section in the discussion. All ‘risk of bias’ tables have been revised (Additional file 4). We have not “track-changed” this additional file, since it is revised comprehensively.

We agree with your comment about samples size and have deleted this from the bias assessment tool.

6. Citing Cochrane methodology. Given that authors have not fully adhered to Cochrane methodology, I would suggest that they adjust their wording to indicate that they used Cochrane methodology as a guide for conducting their review. Examples where Cochrane methodology was not strictly followed include the approach for summary risk of bias assessments, not using GRADE for judging the quality of the evidence, and use of a standardized mean difference where outcomes are measured on the same (as opposed to different) tools.

Reply 6: This has now been rephrased in the manuscript
7. Pg 14, line 24. How did authors interpret 'effect size'? Please provide this information in the methods section.

Reply 7: We agree that this was not clearly defined and we have omitted the wording from the text as this evaluation might be subjective and is related to the outcome of interests.

8. Pg 16, strengths and limitations section. It is not clear to me what the authors feel the limitations of their review are.

Reply 8: Thank you. We have now included a paragraph about limitations in this section.

9. Pg 16, lines 27-28. Cochrane guidance suggests not to provide recommendations in a systematic review; recommendations are for the development of clinical practice guidelines. Authors need to provide conclusions only on what the evidence is.

Reply 9: According to reply 6 about only using the Cochrane methodology as a guideline we would like to keep the section about implications for practice in the manuscript. We have reformulated the section so the difficulty of making conclusions based on this systematic review is clearer.

Minor essential revisions

10. Pg 4, line 5. The authors should make a better distinction that their intent is to update part of the Gagnon and Sandall review, i.e., the evidence regarding small classes. This will give clarity to the reader since the 2007 review also examined individual education.

Reply 10: Our intention was not particularly to update the specific review from Gagnon and Sandall although we acknowledge that some elements are included in both of our reviews. We have used another search strategy, other inclusion- and exclusion criteria as well as an extended bias assessment compared to Gagnon and Sandall. In the ‘strengths and limitations’ section in the discussion, the wording indicated that our study is an update of Gagnon and Sandall’s review. This has now been rephrased.

11. Pg 5. I think the use of the subheadings 'Setting' (lines 5-8) and 'Participants' (lines 8-10) in the Eligibility section would help to structure the information for readers.

We agree and have added the proposed subheadings in the section.
12. Pg 5. For studies that measure the outcome using the same tool, authors should use a 'mean difference' rather than a 'standardized mean difference'. Further, many outcomes were reported by only one study and should be reported with a mean difference.

Thank you. We have now changed the effect measure for continuous outcomes to ‘mean difference’ instead of ‘standardized mean difference’. This has been corrected in the ‘evidence synthesis’ section and in the ‘protocol modifications’ section in the methods section. Also, additional files 6 and 7 have been corrected according to this.

13. pg 7. 'Reporting bias' should be better specified to 'outcome reporting bias', since different types of reporting biases exist.

Thank you for this clarification. We have added this to the manuscript.

14. pg 8, lines 2-3. Authors should concern themselves more with confidence intervals rather than p values. If authors wish to interpret 'borderline significant results', then such interpretation should be provided using confidence interval cutpoints.

Thank you. In the process of going through the analysis again we have decided not to report borderline significant results.

15. Pg 8, lines 8-9. Authors should provide their rationale for including the additional outcomes. They should also be listed in the outcomes section.

Reply 15: We included relationship satisfaction and divorce/separation as secondary outcomes in the review. In the process of going through relevant studies we discovered that these outcomes are also of great relevance as psycho-social dimensions of becoming parents. We have now included a description of this in the manuscript and added them in the outcome section.

16. Pg 8, line 14. Authors should state in their methods section that they contacted authors (of included studies?) to locate additional, potentially relevant studies.

Reply 16: This has now been included in the search strategy section in the methods section.

17. P 8, lines 26-29. This paragraph should follow the one in lines 13-17, since it provides a description of the number of unique studies with multiple reports. Can the authors provide a final tally of unique studies included in the review?
Reply 17: We have now moved the paragraph as proposed and provided the total number of reports (21 papers and one oral presentation).

18. Pg 9, line 7. Provide the missing example for interventions targeting a broader population group.

Reply 18: Thank you. This has now been added.

19. Pg 16, implications for research section. Given the variety of interventions and comparisons observed among included studies, do authors have a suggestion as to where future studies should focus their efforts? Can they comment as to whether general or specific interventions should be the focus? What aspects should be included in such trials? Should there be a focus on the type of comparison?"

Reply 19: Thank you. We have now expanded the ‘implications for research’ section and added our suggestions for the focus of future research.

Reviewer 1:
This is a very comprehensive review of a subject that does need further research. Pregnancy is often thought of as the ‘teachable’ moment where women (and partners) are motivated to do the best for their baby. Yet, we don’t know the best strategies for imparting the information for maximum benefit, or as you point out, whether there are benefits. Intuitively it makes sense that people who have information (regardless of how they get it) should be better prepared, but whether or not that translates into improved outcomes is also a valuable question.

Major Revisions:
The main issue I have with this review is that I think you might be trying to include more than is essential to the review. My understanding is that you wanted to learn more about whether small group antenatal or prenatal classes would make a difference on outcome. My sense when you started the paper was that I was going to be reading about your run-of-the-mill general classes used to educate women about the pregnancy, labour, birth and early parenting processes. Instead there are papers included with very specific content classes. So, in essence you are talking about specific breastfeeding classes, specific psycho-educational antenatal programs for women at high risk of depression, and self-hypnosis intervention classes (which are not really the mainstream type of education). In essence this review is really about whether specialized or general antenatal classes make a difference. My expertise is more on content than methods, so I will ask the editorial group to provide context on this issue.

Reply 1: Thank you so much for your valuable comments and for taking the time to review our manuscript so thoroughly. The main objective of the review is, as you point out, to assess antenatal education in small groups as a method. This has been investigated by comparing with other interventions or no intervention. Current evidence points to the importance of interacting with fellow learners and the learning environment in order to obtain new competencies, and this aspect is our main focus. In the study selection process the
focus has not been on the specific content of class – general or specialized – but on the teaching setting. However, we have now compared groups of interventions based also on the content of the intervention as described according to our answer to the editorial comment no. 3.

In your conclusion you recommend updating this review regularly, but there is also the issue that education will gradually change to web-based models as well as other compressed formats and this is also something that will have to be evaluated. Perhaps it would be worthwhile to spend some time in the discussion talking about how models of education are changing over time and how this might affect a review like this in the future.

We agree that new technology most likely will affect the methods of delivering antenatal education. As stated in our reply above we were interested in the effect, due to group dynamics, of small classes, which we consider may be difficult to obtain in web-based models. However, we agree that new types of education methods should also be evaluated in future reviews when trials have been conducted.

**Minor Revisions:**

1. **Background:** In the first paragraph you state that antenatal education is about providing parents with strategies for dealing with .... In fact however, there is a lot of information imparted on health promotion and risk reduction.

   **Reply 1:** Thank you for pointing to this. We agree and have incorporated this in the section.

2. **Line 10, p. 3** - you say that the type and arrangement of education is debated - I don't know what this means. Similarly in the same line it says antenatal education has been sensitive to opinions and trends and I wasn't sure what you were getting at. In the bottom sentence of the same paragraph, you talk about outcomes relevant to health care providers. However, we should be thinking more about outcomes relevant to women and families.

   **Reply 2:** We have rephrased this section so this is hopefully clearer now.

3. **Lin 11-17 p. 5.** This clearly describes your experimental and control conditions, and supports what you have undertaken to do. I guess my issue is likely with the comparability of these groups and interventions. Education for depressed women would be inherently different than education as an intervention for another group of women without depression. Does the lack of benefit in this case reflect the issue with the small group intervention vs the underlying condition of someone?

   **Reply 3:** In the trials of depression-prevention the participants are screened according to risk of getting depressed. All the high-risk pregnant women are thereafter randomized and therefore the underlying condition should not affect the outcome of antenatal education classes when comparing the two groups. We agree that the effect of an intervention targeted high-risk women on depression will not reflect the
effect you may see if the same trial was conducted in a low-risk population and results are therefore of limited generalizability.

4. I wonder if there should be some words in the discussion based on line 28 (p.5) on whether education will really change obstetric outcomes. All the education in the world won’t change a fetal heart emergency in labour that leads to c-section or the need for induction or the need for operative delivery due to a pelvic issue. While people have made these outcomes as part of the study, are they biologically plausible?

Reply 4: We agree that the scenario you describe will not be prevented by antenatal education in small classes. The authors in some of the studies using obstetric outcomes have argued that a greater sense of control and better coping may prevent anxiety and thereby decrease obstetric interventions. Also, due to the randomized trial design it is assumed that the number of fetal heart rate emergencies are even between the experimental and control condition and will therefore not affect the relative risk.

Although this issue is highly relevant in original papers on trials in this area, we have chosen not to discuss this issue in detail as we try to keep our discussion very concisely related to the specific review aim.

5. The whole issue of BF initiation and duration is hugely dependent on many other factors (other than education). I haven’t looked at the trials, but unless the populations were carefully selected, it isn’t surprising that the results varied.

Comparing trials of mothers, fathers, and general vs specific classes isn’t really comparing a similar intervention.

Reply 5: We agree that the study populations vary greatly and therefore the effect of the different trials are difficult to compare. Moreover, both experimental and control conditions varied greatly between trials which also limits overall conclusions on effect of antenatal education classes. We have revised all analyses and reported results separately for each comparison. We hope that this re-analysis have made the results clearer.

6. Discussion - I thought there might be a bit more of a fullsome discussion about WHY antenatal education results are so difficult. Is there some bias in people who attend small classes vs those who choose other means of education. I’ve already mentioned that education is changing and we need to find the right outcomes to evaluate. It would be important to acknowledge this. What is feasible to measure and what is a good outcome. Is increased knowledge enough or should we be expecting major changes. Is there some other sort of clinical perspective that has similar issues that could be discussed.

As we solely review RCT’s, there should not be large problems of selection bias between the groups. We are unsure about the rest of the discussion points in this point of critique and have therefore not been able to include text specifically addressing these issues.
Reviewer 2:
Thank you for the opportunity to review this paper. The objective of this systematic review is to assess the effects of antenatal education in small groups on obstetric as well as psycho-social outcomes. This is not really a systematic review (as usually these include a meta analysis) – it is more a review of the evidence undertaken in a systematic way.

Reply 1: Thank you very much for taking the time and careful effort to read and comment on our manuscript. We agree that the optimal method for synthesis of results from the involved trials would be a meta-analysis. However, we still believe that this review is a systematic review as it includes all characteristics mentioned in the Cochrane Reviews Handbook. Due to diverse experimental and control conditions between all trials it was unfortunately not possible to perform a meta-analysis.

The authors acknowledge that a review was conducted in 2007 but state that since then more randomized trials have been conducted. It would be useful at this point to have identified which trials have been conducted since the last review.

Reply 2: We have rephrased this section in the introduction so that it is clearer that our review is not an update of the review from Gagnon and Sandall. We have a more narrow focus on the effect of small classes and used another search strategy and other inclusion- and exclusion criteria. Therefore it is not possible to provide the number of conducted trials in the period between the two systematic reviews.

The objective was: to determine the effectiveness of antenatal education in small classes on obstetric and psycho-social outcomes compared to no intervention, treatment as usual or other types of educational programs using randomized trials from Western countries.

Why were only ‘western’ countries included?

Reply 3: We have chosen to include Western countries only (defined as OECD countries) in the systematic review since preparation for birth and parenthood is dependent on culture and contextual factors, such as the organization of the health care system. The purpose of this review is to guide decision making in western countries – therefore the restriction. In the manuscript’s methods section under “eligibility criteria” this is described.

Were quasi randomised trials included? Were all the interventions face to face?
Reply 4: Yes, also quasi-randomized trials were eligible for inclusion. This has now been stated in the eligibility criteria section in the method section. These will score ‘high risk of bias’ in sequence generation and therefore not be given much weight in the presentation of results. All interventions were face-to-face.

There are 4 primary outcomes on pages 5-6 but these are very broad and actually include many more outcomes. Were more specific outcomes determined? For example, what forms of pain relief (was it epidural anaesthesia vs everything else or any form versus nothing?). Which obstetric interventions were included (induction, augmentation, episiotomy, caesarean section etc?).

What were the measures of psychological and social adjustment to parenthood (and this seems like more than one outcome)? Antenatal and postnatal depression and anxiety seems like 4 outcomes not one? The secondary outcomes are also very broad and encompass multiple elements. I would have thought clarity about the outcomes would have been helpful.

Reply 5: We chose beforehand – when writing the protocol for this systematic review - to have broad categories of outcomes due to the diverse aims of antenatal education. Within each category, e.g. pain relief, we have reported all pain relief outcomes in the trial. The outcomes under each broad category are presented in the results section, tables of effect and forest plots.

The authors seem to have included all the trials regardless of risk of bias analysis?

Reply 6: We have chosen to include all trials regardless of risk of bias. We consider this valuable – especially in fields were trials are limited. Many of the included trials were of older date when the tradition for reporting sources of bias in trials was different. Excluding these trials from the review might have resulted in loss of information. By using the ‘risks of bias’ table, and the ‘risk of bias summary’ figure, our judgment about risk of bias is made explicit to the reader. Throughout the Results section, we have made the bias scoring explicit to the reader in order to leave the full information and estimation of the results clear and open to the reader … In the “quality of evidence” section in the discussion we again emphasize the importance of relying on the “low risk of bias” trials.

I was surprised to see trials about group antenatal care included? These trials (eg. Ickovics et al) are not really about education per se – they also provide full antenatal care.

Reply 7: thank you for this comment. In the eligibility criteria we have stated: “The experimental conditions in the included trials must be delivered as an antenatal educational program offered by an educator to groups consisting of more than one individual/couple and less than 20 persons, related to delivery and/or preparation for parenthood”. In the trial by Ickovics they state that “The majority of time is spent with women and clinicians engaging in discussion, education, and skills building to address explicit learning objectives in prenatal care, child birth preparation, and postpartum and parenting roles”. We therefore consider this trial relevant for inclusion in the review.
I was surprised that a meta analysis could not be performed. I would have thought it would have been more useful to be much clearer about the inclusion and exclusion criteria to ensure that there were some similarities of intervention and then do a meta analysis.

Supplementary file 6 presents the RR and 95% CI. This file includes multiple analyses for what seems to be the same outcome (e.g. epidural analgesia). Why was a meta analysis not performed on these outcomes? Presenting the data in this way is not very helpful as it is only really a summary of the different studies but does not seem to move the evidence along at all.

Reply 8: As stated in the protocol for the systematic review we planned to conduct a meta-analysis. However, along the process this turned out not to be feasible due to the heterogeneity of trials. Although we find more than one trial reporting e.g. epidural analgesia as an outcome, all the trials are so diverse in both experimental and control conditions that describing the size of effect using meta-analysis was considered not to be meaningful. This argument is extended in the Cochrane handbook page 245.

We have revised all analyses and reported results separately for each comparison. We hope that this re-analysis has clarified the results and improved the information from our study.

We have now changed the effect measure for continuous outcomes to ‘mean difference’ instead of ‘standardized mean difference’. This has been corrected in the ‘evidence synthesis’ section and the ‘protocol modifications’ section in the methods section. Also, additional files 6 and 7 have been corrected according to this.

The authors state that they “assessed the literature up to 2014 and the results from the trials conducted in the interim period have not altered the conclusion from the earlier review of a largely unknown effect. Also, we have included earlier trials not reported in the systematic review conducted by Gagnon and Sandall – these did not change conclusions.” Given there is no meta analysis presented I am not sure how they can say this?

Reply 9: We have rephrased this section so it is clearer that this systematic review in many ways is different from the one by Gagnon and Sandall.

Level of interest: An article whose findings are important to those with closely related research interests

Quality of written English: Acceptable

Statistical review: No, the manuscript does not need to be seen by a statistician.

Declaration of competing interests: I am part of a group currently updating the previous systematic review on antenatal education (Gagnon AJ, Sandall J: Individual or group antenatal education for childbirth or parenthood, or both. Cochrane database of systematic reviews (Online) 2007:CD002869). Therefore I have a conflict of interest.